

Plenaire vergaderingen Tweede Kamer op beeld bewaard

Archivering van het audiovisuele materiaal van de Tweede Kamer bij Beeld en Geluid

Audiovisuele verslaglegging van politieke vergaderingen wordt steeds belangrijker. Ook in de Tweede Kamer worden van de zittingen van de plenaire vergaderingen en van de meeste commissievergaderingen opnames gemaakt. Deze beelden kunnen door iedereen gezien worden via de app *Debat Direct* of op een later moment bekeken worden, via de website *Debat Gemist*. Het duurzaam opslaan van dit audiovisuele materiaal is echter een hele uitdaging. Het vergt niet alleen specialistische kennis en systemen, maar ook de benodigde opslagcapaciteit.

Tekst **Annelies Cordes en Karolien Verbrugge**

Daarom werken de Tweede Kamer, het Nationaal Archief (NA) en het Nederlands Instituut voor Beeld en Geluid (Beeld en Geluid) sinds 2013 samen aan het duurzaam archiveren van de registraties van de plenaire vergaderingen van de Tweede Kamer bij het Nationaal Archief en Beeld en Geluid daags na creatie. In dit artikel wordt beschreven hoe de drie organisaties het geautomatiseerd proces van het opslaan en toegankelijk maken van de beelden van de plenaire vergaderingen bij Beeld en Geluid hebben ingericht.

De reden dat gekozen is om de archivering van de beelden bij Beeld en Geluid plaats te laten vinden, in plaats van bij het Nationaal Archief, is dat dit instituut ingericht is om dit type archiefmateriaal duurzaam te bewaren en toegankelijk te maken. Mede door de nauwe verbondenheid met het productieproces van de landelijke publieke omroep is het digitale archief van Beeld en Geluid op basis van geavanceerde technologie van aanvang af ingericht op instroom, opslag, beschikbaarstelling en professioneel hergebruik.

De digitale infrastructuur is de afgelopen jaren verrijkt met voorzieningen voor toekomstbestendige digitale preservatie. Aan de hand van de hiervoor ontwikkelde processen en technieken voor kwaliteitscontrole, lifecyclemanagement en gegevensmodellering kan de authenticiteit, de integriteit

en de afspeelbaarheid van alle ingestroomde audiovisuele content blijvend worden gegarandeerd. Beeld en Geluid is gecertificeerd als 'Trustworthy Digital Repository'.

Pilots

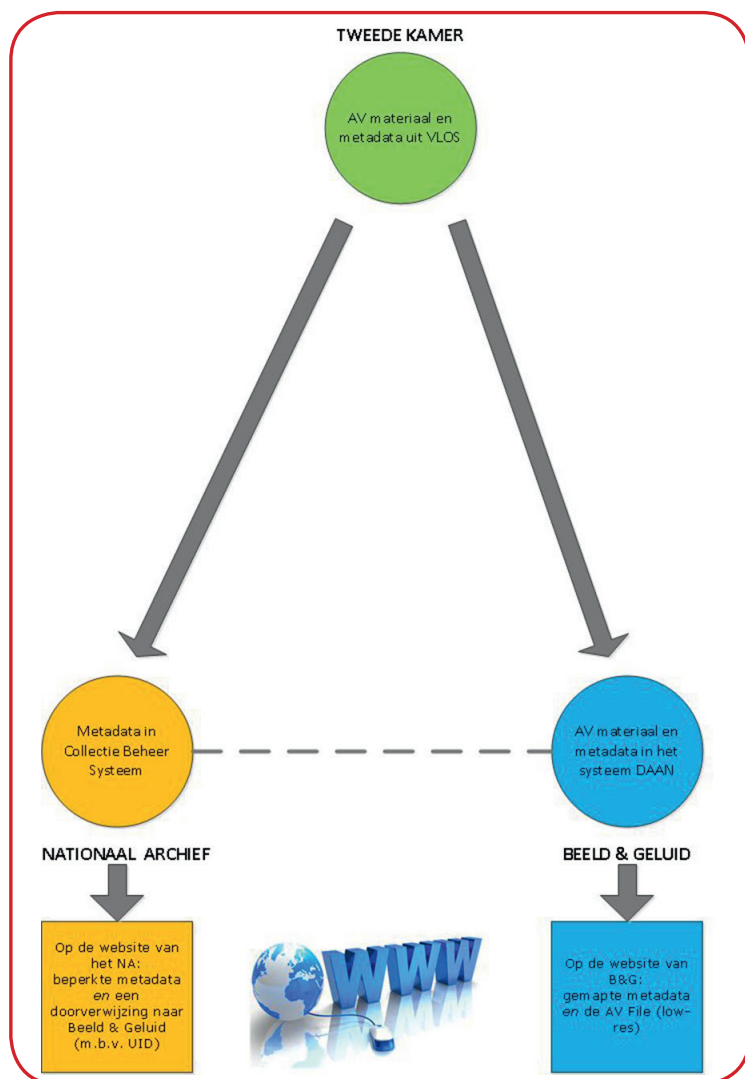
Sinds 2013 hebben de Tweede Kamer, het Nationaal Archief en Beeld en Geluid drie pilots uitgevoerd waarvan er inmiddels twee zijn afgerond. Het doel van de drie pilots is om een volledig geautomatiseerd proces van archivering van av-materiaal en metadata in te richten. In de eerste pilot in 2013-2014 is door de drie partijen gekeken welke stappen genomen moesten worden om tot een automatische instroom van hoge resolutie-beelden en metadata te komen. In de tweede pilot is de infrastructuur voor de au-

Sinds 2015 worden de beelden van de vergaderingen van de Tweede Kamer daags na creatie doorgestuurd naar Beeld en Geluid

tomatische instroom en archivering van de beelden gerealiseerd. Sinds 2015 worden, als resultaat hiervan, de beelden van de vergaderingen van de Tweede Kamer daags na creatie doorgestuurd naar Beeld en Geluid.

In de derde pilot, die in 2017 van start is gegaan, staan twee zaken centraal: (1) de archivering van de beelden in het Digitaal Audiovisueel Archief van Nederland (DAAN) – het nieuwe Media Asset Management-systeem¹ (MAM) van Beeld en Geluid – en (2) het verder toegankelijk maken van de vergaderingen bij Beeld en Geluid en het Nationaal Archief. De drie instituten willen in de loop van 2019 het hele proces van instroom, archivering en publicatie operationeel hebben.

¹ Een Media Asset Management-systeem is een systeem dat digitale media assets opslaat, beheert en op eenvoudige wijze toegankelijk maakt. Denk bij 'media assets' aan AV-materiaal, foto's of digitale gidsen die de benodigde metadata bevatten.



AV-files

De hoge resolutie-beelden van de plenaire vergaderingen van de Tweede Kamer worden opgeslagen bij een externe leverancier die de beelden een aantal dagen bewaart. Binnen deze periode moeten deze hoge resolutie-beelden doorgestuurd worden naar en gearchiveerd worden bij Beeld en Geluid. De beelden worden aan Beeld en Geluid aangeleverd als een MXF-bestand. Beeld en Geluid hanteert de open MXF-standaard om duurzame opslag van av-materiaal te garanderen. Deze standaard is naast open ook gangbaar en wordt onderhouden en ondersteund door grote fabrikanten. Een ander voordeel van het hanteren van deze standaard bij Beeld en Geluid is dat het formaat mediapark-breed wordt gebruikt en het materiaal daardoor direct voor uitzending geschikt is.

Voordat de beelden worden doorgezonden naar Beeld en Geluid, levert de Tweede Kamer een kleine set metadata aan de leverancier aan. Met deze metadata kan in het beeldmateriaal de start en het einde van de vergadering worden bepaald. De 'geknipte' beelden worden vervolgens door de leverancier doorgestuurd naar Beeld en Geluid via een beveiligde ftp-verbinding.

Metadata

De metadata die de Tweede Kamer aanlevert aan Beeld en Geluid worden gebruikt om de av-registraties te ontsluiten,

zodat het beeldmateriaal optimaal toegankelijk is. De metadata zijn afkomstig uit VLOS, het verslagleggingsondersteunend systeem van de Tweede Kamer. Dit systeem wordt gebruikt ter ondersteuning van de verslaglegging van de vergaderingen door de Dienst Verslag en Redactie. Medewerkers markeren alle gebeurtenissen die tijdens de vergadering voorvallen. In dit xml-bestand staat onder andere wie, wanneer over welk onderwerp heeft gesproken.

In eerste instantie wordt het ongecorrigeerd verslag van de vergadering aangeleverd aan Beeld en Geluid. Dit verslag is bijna direct na afloop van de plenaire vergadering beschikbaar, maar is nog in concept. Er kunnen dus nog wijzigingen komen in de metadata. Dit verslag (VLOS-xml) stuurt de Tweede Kamer tezamen met de av-beelden en de checksums² aan Beeld en Geluid.

De ontvangen metadata in de VLOS-xml worden zorgvuldig gemapt/vertaald naar de velden in DAAN, het MAM-systeem van Beeld en Geluid. Deze metadata maken het mogelijk om de inhoud van de gemiddeld acht uur durende plenaire vergaderingen op verschillende niveaus vindbaar en doorzoekbaar te maken voor het publiek.

Zodra bij de Tweede Kamer het gecorrigeerde verslag gereed is, wordt dit aangeleverd bij Beeld en Geluid. De nieuwe VLOS-xml van het gecorrigeerde verslag overschrijft de metadata van het ongecorrigeerde verslag in het digitale archief.

Instroom workflow in digitaal archief

Vanuit de Tweede Kamer wordt een complete importset aangeleverd bij Beeld en Geluid. Deze set bestaat uit drie bestanden: de Tweede Kamer Plenaire Vergaderfile (av-file), de VLOS-xml en een checksum berekend over de av-file.

Het laatste bestand dat compleet is geüpload naar de ftp-server van Beeld en Geluid is voor DAAN de trigger om de import-workflow te starten. Deze import-workflow omvat diverse controles, waaronder de checksum controle, een (technische) kwaliteitsanalyse van de av-file en metadata-xml en een volledigheidscntrole op de av-file en de metadata-xml.

Wanneer de controles geen fouten opleveren, worden de Vergaderfile en de VLOS-xml gezamenlijk als 'archiefobject' opgeslagen in het digitale archief van Beeld en Geluid.

Persistent Identifier

Zodra de files zijn ingestroomd in het digitale archief van Beeld en Geluid, start vervolgens de *Persistent Identifier workflow*. Tijdens deze workflow wordt een Persistent Identifier³ (PID) toegekend aan het archiefobject. Dit is

² Checksums (controlesommen) worden gebruikt om de integriteit van een bestand te kunnen garanderen nadat het is verzonden van de ene (digitale) opslagplaats naar de andere. De checksum wordt berekend met behulp van een algoritme, de uitkomst hiervan wordt samen met de file verstuurd. Om de integriteit van het bestand te verifiëren, berekent de ontvanger de checksum op de ontvangen file opnieuw. Als de twee checksums overeenkomen is het bestand ongewijzigd.

³ Een *persistent identifier* is een unieke identificatiecode van een digitaal object die onafhankelijk is van de bewaarlocatie. Door een unieke identificatie aan de informatie, in plaats van aan het webadres, toe te kennen, blijven gegevens duurzaam toegankelijk, ook als de informatie van plaats verandert.



De Tweede Kamer stemt tijdens de vergadering op 25 april 2019. Bron: Tweede Kamer

belangrijk voor de duurzame opslag van digitale objecten omdat de eindgebruiker een digitaal object met een Persistent Identifier altijd kan terugvinden en tot de bron kan herleiden. Zo worden 'dode' links voorkomen en zullen referenties naar de vergaderingen van de Tweede Kamer vanuit Nationaal Archief naar Beeld en Geluid blijven bestaan.

Via de Persistent Identifier kan de eindgebruiker een digitaal object altijd terugvinden en tot de bron herleiden

De Tweede Kamer vraagt de PID na archivering op bij Beeld en Geluid via een API-interface en weet op die manier dat het archiefobject van een plenaire vergadering goed is opgeslagen. De PID wordt ook door de Tweede Kamer bewaard voor doorgifte naar het Nationaal Archief.

Overbrenging

De Tweede Kamer hoort haar data en metadata over te brengen naar het Nationaal Archief. Het Nationaal Archief heeft er in deze pilot voor gekozen om het av-materiaal te laten opslaan bij Beeld en Geluid, als specialist op het gebied van archivering en beschikbaar stellen van media. Er is nu nog sprake van uitplaatsing van het av-materiaal van de Tweede Kamer, maar in de toekomst zal het av-materiaal inclusief metadata worden overgebracht. Dit gebeurt nog steeds naar het Nationaal Archief, maar de feitelijke opslag, ontsluiting en publicatie van het av-materiaal zal plaatsvinden bij Beeld en Geluid.

De Tweede Kamer levert ten behoeve van de overbrenging periodiek een aantal metadata aan het Nationaal Archief aan, inclusief de PID's. Dit is om te duiden welke archiefobjecten worden overgebracht naar het Nationaal Archief. Deze metadata worden toegevoegd aan de inven-

taris van het av-materiaal van de Tweede Kamer. Met het aanleveren van deze metadata wordt in praktische zin de overbrenging geregeld.

Wie in de toekomst via de website van het Nationaal Archief op zoek gaat naar beelden van de plenaire vergaderingen van de Tweede Kamer, wordt via de PID's in de inventaris doorverwezen naar de portal van Beeld en Geluid. Daar kunnen de beelden door iedereen bekeken worden.

Toekomst

Binnen de scope van de huidige pilot valt ook het met terugwerkende kracht aanleveren aan Beeld en Geluid van de xml-verslagen voor de ontsluiting van het beeldmateriaal dat al bij Beeld en Geluid staat. Wanneer het proces volledig operationeel is en dagelijks de data en metadata van de plenaire vergaderingen gearchiveerd worden bij Beeld en Geluid, is de intentie om ook op deze wijze het av-materiaal van de commissievergaderingen van de Tweede Kamer duurzaam op te slaan en toegankelijk te maken. ●



Annelies Cordes

Coördinator Instroom bij het NIBG



Karolien Verbrugge

Archivaris bij de Tweede Kamer